Solutions to Quantitative Researcher Puzzle

Jason Barrella

November 16, 2021

Final Answers

1. **2/3**
2. 1/4 or 3/4
3. 1/6 or 5/6

Explanation of Parts 1 & 2

The approach taken to solve this problem was first attempting to find a solution by inspection and then to confirm it numerically. In this problem, there are three contestants who each have to select a unique real number x in the range [0, 1], that is $x \in \mathbb{R} \mid 0 \leq x \leq 1$. After all three contestants have made their selection, a random real number in the same range is drawn from a uniform distribution. This means that all numbers in the range are equally likely to be drawn. The contestant who's number is closest to the randomly drawn number is the winner. The problem revolves around computing probabilities for each contestant to win given their individual selections. Since the random number is drawn from a uniform distribution and the range is already normalized to 1, the probability for any contestant to win is simply equal to the length of the range of numbers that they are nearest to.

Let us denote A's choice as a, B's choice as b, and likewise, C's choice as c. We will also use $P(A)$, $P(B)$, and $P(C)$ to denote the probabilities for A, B, and C to win respectively. Contestant A chooses 0 so we have $a = 0$ for part 1. The optimal choice for B is the choice that results in them being nearest to the largest possible range of numbers under the assumption that C will make the optimal selection when their turn arrives.

Let us first begin by identifying some fundamental principles of the game. Firstly, there is a symmetry about 0.5, in other words, an opening move of, for example, 0.8 is mathematically identical to an opening move of $1 - 0.8 = 0.2$ and will result in the same outcome. This is for the same reason as the fact that the game would not be changed if the board were shifted from [0, 1] to [1, 2], for example. What matters in the game are the lengths between the points involved which are a, b, c, and the endpoints. Lengths are invariant under translations and rotations. Using the range [0, 1] is however useful as it gives us the probabilities directly. With this in mind, for the opening move of A, we need only consider the range [0, 0.5] without any loss of generality. We can then include the equivalent move $a' = 1 - a$ later.

It is also useful at this stage to point out a mathematical feature of the game which is that, for any $x < y < z$ where x, y, and z are each one of a, b, c in no particular order,

$$P(Y) = \frac{(y - x) + (z - y)}{2} \tag{1}$$

$$= \frac{z - x}{2} \tag{2}$$

which is independent of y. In other words, any move made by one player anywhere in the region between two other players awards that player with equal probabilities of winning.

In order to determine the best move for A, we need to consider the best strategies for B given a particular a. After A has played, B must choose some value between a and one of the outside endpoints, either 0 or 1. It is clear to see that space between b and either of the endpoints is twice as valuable as space between b and a since the latter is shared between the two players. With this is mind, B's objective becomes to maximize the amount of retained outside space, that is, space between itself and one of the endpoints. Further thought reveals that B would never play below A given $0 \leq a < 0.5$ since, by playing above a, B could always control more space including outside space. Broadly speaking, B should play a move such that C is forced to play below B so that B achieves this objective. This can be done in one of two ways, the choice of which depends on a.

Small a

For $a \leq 1/4$, B's optimal move is to play 2/3 of the remaining space up the board, that is

$$b = 2(1-a)/3 + a. \tag{3}$$

This divides the remaining board space above a into two regions of slightly less than equal value for C. The goal of this is to force C to play between a and b, leaving B to occupy more outside space and to profit on the average of C's random draws from the infinite optimal moves that is has in the region between a and b as shown in eq. 2. An even better outcome for B would be to force C to play below a, but for $a \leq 1/4$, this is not possible as will be shown below. Let us examine the optimal move for C in each of the three regions with $b = 2(1-a)/3 + a$.

Above B
If C were to play above B, the best move would clearly be $c = b + \epsilon$ where $\epsilon \ll 1$ is used to denote the smallest possible delta in the game. This would award C with

$$\begin{aligned} P(C) &= 1 - b - \epsilon + \epsilon/2 \qquad (4) \\ &= 1 - b - \epsilon/2 \qquad (5) \end{aligned}$$

Between A and B
If C were to play between A and B, C is awarded $P(C) = (b-a)/2$. However, since we know that $b - a = 2(1-b)$ because B played 2/3 of the remaining board space up the board, we can then show that

$$\begin{aligned} P(C) &= (b-a)/2 \qquad (6) \\ &= 1 - b. \qquad (7) \end{aligned}$$

Since $1 - b - \epsilon/2 < 1 - b$, we can conclude that C's optimal choice will not be to play above B. One can also see that if B played anywhere below their optimal move, C would have played above b and B would have been in a much worse position where they would occupy no significant outside space and would share all inside space with A which would be strictly less than half the outside space.

Below A
Lastly, by playing $c = a - \epsilon$, C will similarly be awarded with

$$P(C) = a - \epsilon/2 \tag{8}$$

Therefore, C will not play below A if $(b-a)/2 > a - \epsilon/2$. Using $(b-a)/2 = 1/3(1-a)$ we can rewrite this condition as

$$a = (b-a)/2 \tag{9}$$
$$= 1/3(1-a) \tag{10}$$
$$= 1/4. \tag{11}$$

Which gives us back the maximum value of a for this strategy to work. In this scenario, B is awarded with a winning probability of $P(B) = (1-b) + (b-c)/2$. Since C will randomly draw from the infinite optimal moves in the region between a and b, we have an expectation value for c of $E[c] = (b+a)/2$ which gives an average probability over successive games for B to win of

$$P(B) = (1-b) + (b-c)/2 \tag{12}$$
$$= (1-b) + (b - [(b+a)/2])/2 \tag{13}$$
$$= (1-b) + (b/4 - a/4) \tag{14}$$
$$= 1 - 3/4b - a/4 \tag{15}$$
$$= 1/2 - a/2 \tag{16}$$

using eq. 3. For $a = 0$, this equation yields $P(B) = 1/2$. This was confirmed numerically with Python. The results are shown below in fig. 1.

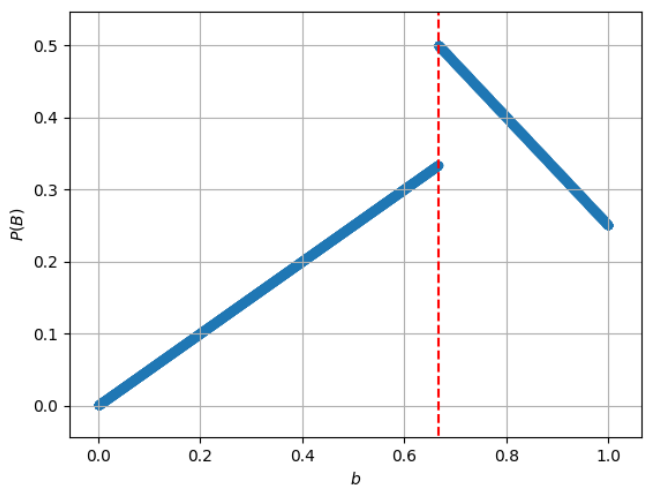

Figure 1: Average over successive games of the probability for player B to win given varying choices of b. A's move of $a = 0$ is fixed. The red dashed line is drawn at $b = 2/3$. A discontinuity is observed here as C's strategy changes from playing $c = b + \epsilon$ for values of $b < 2(1-a)/3 + a$ to playing anywhere in the region between a and b for values of $b \geq 2(1-a)/3 + a$.

Large a

For $a > 1/4$, B employs a different strategy. This is because 1/3 of the remaining space on the board is less than a. In other words, B does not need to play at exactly $b = 2(1-a)/3 + a$ anymore since C will play $c = a - \epsilon$ for smaller values of b which, in turn, will give B more outside board-space and therefore larger $P(B)$. B's optimal choice now is to play at

$$b = (1-a) + \epsilon. \tag{17}$$

Let us again examine the best possible move for C in each of the three available regions in such a scenario.

Above B

C's best choice in this region will, again, be to play $c = b + \epsilon$ to gain as much outside space as possible. This will award C with $P(C) = 1 - b - \epsilon/2$ as seen in eq. 5. With $b = (1 - a) + \epsilon$, we get

$$\begin{aligned} P(C) &= (1 - [(1 - a) + \epsilon] - \epsilon/2 &(18)\\ &= a - 3\epsilon/2 &(19) \end{aligned}$$

Between A and B

For any choice of $a < c < b$, we have $P(C) = (b - a)/2$ as before. Now in the case of $a > 1/4$ and $b = (1 - a) + \epsilon$ we obtain

$$\begin{aligned} P(C) &= (b - a)/2 &(20)\\ &= (1 - a + \epsilon - a)/2 &(21)\\ &= 1/2 - a + \epsilon/2 &(22)\\ &< 1/4 &(23) \end{aligned}$$

With $a \geq 1/4 + \epsilon$.

Below A

This is the same as in the previous case for small a where we had,

$$\begin{aligned} P(C) &= a - \epsilon/2 &(24)\\ &> 1/4 &(25) \end{aligned}$$

Again, using $a \geq 1/4 + \epsilon$ and also $c = a - \epsilon$. This is clearly larger than $P(C)$ in the region above B as seen in eq. 19. It is also larger than $P(C)$ from eq. 23 when considering the definition of ϵ as being the smallest allowable difference between two moves.

Since $P(C)$ is larger in the region below A, the above demonstrates that C will play below A and B will be awarded with the maximum probability of winning possible for these particular choices of a and b.

This also shows that $a > 1/4$ is not the optimal move for A. We have seen that A cannot force C to play above b since B always has a move that forces C to play below b. This means that the optimal move for A is the one that retains the maximum amount of outside space and forces C to play in between a and b. This is exactly the upper limit of small a which is 1/4. Therefore, the best opening move for A is $a = 1/4$. This awards A with a chance to win of

$$\begin{aligned} P(A)_{a \leq 1/4} &= a + (c - a)/2 &(26)\\ &= a + ([(b - a)/2 + a] - a)/2 &(27)\\ &= a + b/4 - a/4 &(28)\\ &= a + 1/6 - a/6 + a/4 - a/4 &(29)\\ &= 1/6 + 5a/6 &(30) \end{aligned}$$

using eq. 3 for b. This gives A a maximum chance of winning of $P(A) = 3/8$ at $a = 1/4$. This was also confirmed numerically in Python. The results are shown below in figure 2.

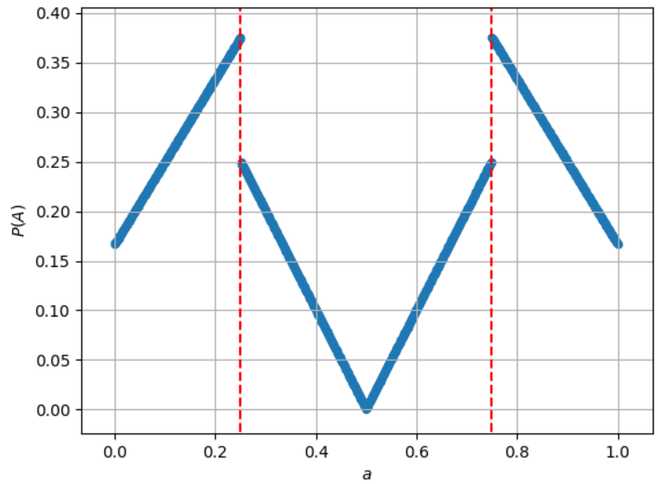

Figure 2: Average over successive games of the probability for player A to win given varying choices of a. The red dashed lines are drawn at $a = 1/4$ and $a = 3/4$. At these points, discontinuities are, again, observed where player C adjusts their strategy from playing between a and b to playing between a and the edge of the board. Symmetry is observed around $a = 0.5$.

Numerical Results for Part 3

For this question, a fourth player is introduced whom shall be called player D. No analytical solution was attempted, instead a simulation was written in Python that found the optimal values for each of b, c, d given each possible opening choice of a. Once these are known, the probability for A to win can be computed. The results from the simulation are shown below.

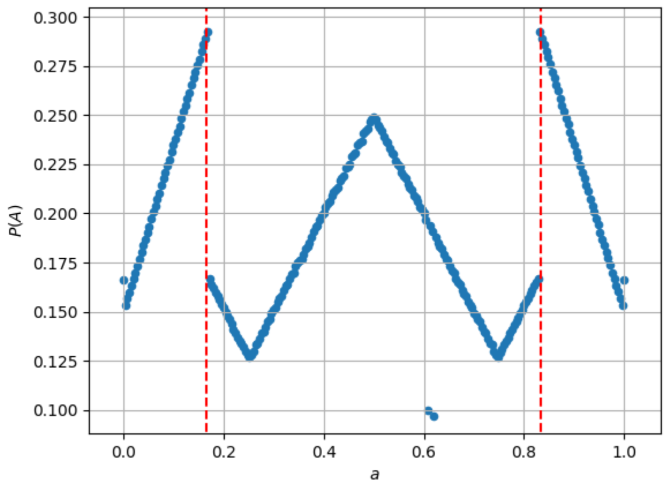

Figure 3: Average over successive games of the probability for player A to win given varying choices of a in the case of four players. The red dashed lines are drawn at $a = 1/6$ and $a = 5/6$. $p = 250$ points were tested in the range [0, 1].

The optimal move for A is found to be $a = 1/6$ or $5/6$. A choice of $a = 1/6$ results in B and C playing $b = 5/6$ and $c = 1/2$ respectively. B plays $b = 1/6$ in the case of $a = 5/6$ leading to the same outcome for C. D is then left to draw from infinite optimal moves in the range $c = (1/6, 5/6)$. In this scenario, A achieves the maximum chance to win of

$$
\begin{aligned}
P(A) &= 1/6 + [3(1/2 - 1/6)/4]/2 && (31)\\
&= 7/24 && (32)\\
&= 0.291\overline{6} && (33)
\end{aligned}
$$

Some outliers are also observed. These were found to be artifacts caused by the limited resolution of the discrete points that were tested in the range [0, 1] for each player. $P(A)$ was found to approach 7/24 over consecutive simulations with varying p. The program made use of three nested for loops meaning the complexity increases as p^3 making large p computationally expensive. More work could be put into optimizing the code for better performance if larger p needed to be tested.

Speculative Generalization

Comparing solutions for varying numbers of players, a trend begins reveal itself. For $n = 2$ players, the optimal choice for A is $a = 1/2$, for $n = 3$ players, the optimal choice was found to be $a = 1/4$, and for $n = 4$ players, the optimal choice has now been discovered to be $a = 1/6$. This progression suggests a general solution for optimal a given n of $a = \frac{1}{2(n-1)}$. The symmetric move of $1 - a$ may also be played for any n. Observing the similarities in the structures of fig. 2 and 3 invites further speculation. It seems plausible to think that, for all n, there will be positively sloped lines pointing inward on each outside edge that both reach the global maximum at $a = \frac{1}{2(n-1)}$ and $a = 1 - \frac{1}{2(n-1)}$ as well as a central peak/trough that oscillates between pointing up and pointing down for even and odd values of n respectively. Well this hypothesis may seem possible, further data is needed to support it.